

SEQUENCE LISTING

(I) GENERAL INFORMATION

(i) APPLICANT: Goli, Surya K.  
Hillman, Jennifer L.

(ii) TITLE OF THE INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: INCYTE GENOMICS, INC.  
(B) STREET: 3160 Porter Drive  
(C) CITY: Palo Alto  
(D) STATE: CA  
(E) COUNTRY: US  
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/309,320  
(B) FILING DATE: May 11, 1999  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09 309,320  
(B) FILING DATE: May 11, 1999  
  
(A) APPLICATION NUMBER: 09 096,571  
(B) FILING DATE: June 12, 1998  
  
(A) APPLICATION NUMBER: 08 756,771  
(B) FILING DATE: November 26, 1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.  
(B) REGISTRATION NUMBER: 36,749  
(C) REFERENCE DOCKET NUMBER: PP-3160-3

(ix) TELECOMMUNICATION INFO/MAIL IN

(A) TELEPHONE: 415-962-0366  
(B) TELEFAX: 415-845-4166  
(C) TELEX:

(x) INFORMATION FOR SEQ ID NO:1:

(xi) SEQUENCE CHARACTERISTICS:

PF-1200-100

### (vii) IMMEDIATE SOURCE

(A) LIBRARY: Consensus  
 (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Ala	Ala	Arg	Pro	Lys	Leu	His	Tyr	Pro	Asn	Gly	Arg	Ala	Arg	Met
1					5				10					15	
Glu	Ser	Val	Arg	Trp	Val	Leu	Ala	Ala	Gly	Val	Glu	Phe	Asp	Glu	
					20				25					30	
Glu	Phe	Leu	Glu	Thr	Lys	Glu	Gln	Leu	Tyr	Lys	Leu	Gln	Asp	Gly	Asn
					35				40					45	
His	Leu	Leu	Phe	Gln	Gln	Val	Pro	Met	Val	Glu	Ile	Asp	Gly	Met	Lys
					50				55					60	
Leu	Val	Gln	Thr	Arg	Ser	Ile	Leu	His	Tyr	Ile	Ala	Asp	Lys	His	Asn
					65				70					75	
Leu	Phe	Gly	Lys	Asn	Leu	Lys	Glu	Arg	Thr	Leu	Ile	Asp	Met	Tyr	Val
					85				90					95	
Glu	Gly	Thr	Ile	Asp	Leu	Leu	Glu	Leu	Ile	Met	His	Pro	Phe	Leu	
					100				105					110	
Lys	Pro	Asp	Asp	Gln	Gln	Lys	Glu	Val	Val	Asn	Met	Ala	Gln	Lys	Ala
					115				120					125	
Ile	Ile	Arg	Tyr	Phe	Pro	Val	Phe	Glu	Lys	Ile	Leu	Arg	Gly	His	Gly
					130				135					140	
Gln	Ser	Phe	Leu	Val	Gly	Asn	Gln	Leu	Ser	Leu	Ala	Asp	Val	Ile	Leu
					145				150					155	
Leu	Gln	Thr	Ile	Leu	Ala	Leu	Glu	Glu	Lys	Ile	Pro	Asn	Ile	Leu	Ser
					165				170					175	
Ala	Phe	Pro	Phe	Leu	Gln	Glu	Tyr	Thr	Val	Lys	Leu	Ser	Asn	Ile	Pro
					180				185					190	
Thr	Ile	Lys	Arg	Phe	Leu	Glu	Pro	Gly	Ser	Lys	Lys	Pro	Pro	Pro	
					195				200					205	
Asp	Glu	Ile	Tyr	Val	Arg	Thr	Val	Tyr	Asn	Ile	Phe	Arg	Pro		
					210				215					220	

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: ~0.1 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) IMMEDIATE SOURCE:

(A) LIBRARY: Consensus  
(B) CLONE: Consensus

ATGGGAAACAA CGGCGAAAGCTT CGACTTAATTC AATGGAAAGCTT GATTTGAACTA  
 TGGGTATTAG CTGCGCGCGC AGTCCGAGTTT CATGAAGAAT TTGTCGAAAAC  
 TTGTACAAGT TGCAGGATGC TAACCACCTG CTGTTCCAAC AAGTGCCCCT  
 GACGGCATGA AGTTGGTACA CACCCGAAAGC ATTCTCCACT ACATACCGAGA  
 CTCTTTGGCA AGAACCTCAA CGAGAGAACC CTGATTCACA TGTACCTGGA CGGGCACACTG  
 GATCTGTCG AAATGGTTAA CATGCAATCTT TTAAAGAAC GAGATGATNA GAAAAACAAA  
 GATGAAATA TGGTCAAGA CGGAAATATTA AATATATTC TGGCGTTGA AAGAACTTA  
 AGGGCTTACG GATGAAACGTT TGTCTGGCTT ATGAGCTTA GATTTGAACTA  
 TGGGTATTAG CTGCGCGCGC AGTCCGAGTTT CATGAAGAAT TTGTCGAAAAC  
 TTGTACAAGT TGCAGGATGC TAACCACCTG CTGTTCCAAC AAGTGCCCCT  
 GACGGCATGA AGTTGGTACA CACCCGAAAGC ATTCTCCACT ACATACCGAGA  
 CTCTTTGGCA AGAACCTCAA CGAGAGAACC CTGATTCACA TGTACCTGGA CGGGCACACTG  
 GATCTGTCG AAATGGTTAA CATGCAATCTT TTAAAGAAC GAGATGATNA GAAAAACAAA  
 GATGAAATA TGGTCAAGA CGGAAATATTA AATATATTC TGGCGTTGA AAGAACTTA

TACAGTCATG TCTTAATGGGA TCCCAGCTCT GTCATGGTGC TATCTATGTA TTAAGTTGGC  
TCCTAAGTTG GGTTTTTTGT

780  
300

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 825605

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Ala	Glu	Lys	Pro	Lys	Ieu	His	Tyr	Ser	Asn	Thr	Arg	Gly	Arg	Met
1															15
Glu	Ser	Ile	Arg	Trp	Ieu	Ieu	Ala	Ala	Ala	Gly	Val	Glu	Phe	Glu	Glu
															30
Lys	Phe	Ile	Lys	Ser	Ala	Glu	Asp	Leu	Asp	Lys	Leu	Arg	Asn	Asp	Gly
															45
Tyr	Leu	Met	Phe	Gln	Gln	Val	Pro	Met	Val	Glu	Ile	Asp	Gly	Met	Lys
															50
															60
Leu	Val	Gln	Thr	Arg	Ala	Ile	Leu	Asn	Tyr	Ile	Ala	Ser	Lys	Tyr	Asn
															80
Leu	Tyr	Gly	Lys	Asp	Ile	Lys	Glu	Lys	Ala	Leu	Ile	Asp	Met	Tyr	Ile
															95
Glu	Gly	Ile	Ala	Asp	Leu	Gly	Glu	Met	Ile	Leu	Leu	Leu	Pro	Phe	Thr
															110
Gln	Pro	Glu	Glu	Gln	Asp	Ala	Lys	Leu	Ala	Leu	Ile	Gln	Glu	Lys	Thr
															125
Lys	Asn	Arg	Tyr	Phe	Pro	Ala	Phe	Glu	Lys	Val	Leu	Lys	Ser	His	Gly
															130
															140
Gln	Asp	Tyr	Leu	Val	Gly	Asn	Lys	Leu	Ser	Arg	Ala	Asp	Ile	His	Leu
															150
															160
145															
Val	Glu	Leu	Leu	Tyr	Tyr	Val	Glu	Glu	Leu	Asp	Ser	Ser	Leu	Ile	Ser
															175
155															
Ser	Phe	Pro	Leu	Leu	Lys	Ala	Leu	Lys	Thr	Arg	Ile	Ser	Asn	Leu	Pro
															180
															190
165															
Thr	Val	Lys	Lys	Phe	Leu	Gln	Pro	Gly	Ser	Pro	Arg	Lys	Pro	Pro	Met
															195
															205
195															
Asp	Glu	Lys	Ser	Leu	Glu	Glu	Ser	Arg	Lys	Ile	Phe	Arg	Phe		
															210
															215
															220

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 159141

1	5	10	15
Glu Ser Thr Arg Trp Leu Leu Ala Ala Ala Gly Val Glu Phe Glu Glu			
20	25	30	
Lys Phe Ile Lys Ser Ala Glu Asp Leu Asp Lys Leu Arg Asn Asp Gly			
35	40	45	
Tyr Leu Met Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys			
50	55	60	
Leu Val Gln Thr Arg Ala Ile Leu Asn Tyr Ile Ala Ser Lys Tyr Asn			
65	70	75	80
Leu Tyr Gly Lys Asp Ile Lys Glu Arg Ala Leu Ile Asp Met Tyr Ile			
85	90	95	
Glu Gly Ile Ala Asp Leu Gly Glu Met Ile Leu Leu Pro Val Cys			
100	105	110	
Pro Pro Glu Glu Lys Asp Ala Lys Leu Ala Leu Ile Lys Glu Lys Ile			
115	120	125	
Lys Asn Arg Tyr Phe Pro Ala Phe Glu Lys Val Leu Lys Ser His Gly			
130	135	140	
Gln Asp Tyr Leu Val Gly Asn Lys Leu Ser Arg Ala Asp Ile His Leu			
145	150	155	160
Val Glu Leu Leu Tyr Val Glu Glu Leu Asp Ser Ser Leu Ile Ser			
165	170	175	
Ser Phe Pro Leu Leu Lys Ala Leu Lys Thr Arg Ile Ser Asn Leu Pro			
180	185	190	
Thr Val Lys Lys Phe Leu Gln Pro Gly Ser Pro Arg Lys Pro Pro Met			
195	200	205	
Asp Glu Lys Ser Leu Gln Glu Ala Arg Lys Ile Phe Arg Phe			
210	215	220	

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 223 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 193710

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ala Ala Lys Pro Lys Leu Tyr Tyr Phe Asn Gly Arg Gly Arg Met			
1	5	10	15
Glu Ser Ile Arg Trp Leu Leu Ala Ala Ala Gly Val Glu Phe Glu Glu			
20	25	30	
Gly Phe Leu Gln Ile Arg Gln Gln Tyr Gln Lys Met Glu Lys Asp Gly			
35	40	45	
Ile Ile Ile Phe Gly Glu Ile Ile Ile Val Ile Ile Asp Gly Met Met			
50	55	60	
Leu Thr Gln Ile Arg Ala Ile Leu Ser Tyr Leu Ala Ala Lys Tyr Asn			
65	70	75	80
Leu Tyr Gly Lys Asp Leu Lys Glu Arg Val Arg Ile Asp Met Tyr Ala			
85	90	95	
Asp Gly Thr Gln Asp Ile Met Met Ile Ala Val Ala Pro Phe Iys			
100	105	110	
Ile Pro Lys Gln Iys Gln Gln Ser Tyr Asn Leu Ile Ile Ile Asp Ala			

PF 3162 b DIV

145	150	155	160
Leu Glu Ala Ile Leu Met Val Glu Glu Leu Ser Ala Pro Val Leu Ser			
165	170	175	
Asp Phe Pro Leu Leu Gln Ala Phe Lys Thr Arg Ile Ser Asn Ile Pro			
180	185	190	
Thr Ile Lys Lys Phe Leu Gln Pro Gly Ser Gln Arg Lys Pro Pro Pro			
195	200	205	
Asp Gly Pro Tyr Val Glu Val Val Arg Ile Val Leu Lys Phe			
210	215	220	